



Evolutionary studies of few species belonging to Leguminosae family based on RBCL gene

Sagar Patel^{1✉}, Hetalkumar Panchal²

1. Research scholar, G. H. Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujarat-388120, India
2. Assistant Professor, G. H. Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujarat-388120, India

✉**Correspondence to:** Research scholar, G. H. Patel Post Graduate Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidyanagar, Gujarat-388120, India; Mail: sgr308@gmail.com

Publication History

Received: 16 November 2013

Accepted: 20 December 2013

Published: 1 January 2014

Citation


Sagar Patel, Hetalkumar Panchal. Evolutionary studies of few species belonging to Leguminosae family based on RBCL gene. *Discovery*, 2014, 9(22), 38-50

Publication License



© The Author(s) 2014. Open Access. This article is licensed under a [Creative Commons Attribution License 4.0 \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

General Note

 Article is recommended to print as color digital version in recycled paper.

ABSTRACT

Evolutionary study is necessary to find out species' relationship to other species. In this paper, rbcL gene sequences data are considered for analysis. It has fairly conservative rate of evolution. The function of the rbcL gene is to code for the large subunit of ribulose 1, 5 bisphosphate carboxylase/oxygenase (RUBISCO or RuBPCase). The sequence data of the rbcL gene are widely used in the reconstruction of phylogenies throughout the seed plants. Leguminosae family is one of the largest families that contain thousands of species of Plants, Herbs, Shrubs and Trees worldwide. Legumes are used as crops, forages and green manures; they also synthesize a wide range of natural products such as flavours, drugs, poisons and dyes. Legumes have the potential to play an increasing role in alleviating these threats owing to their ability to produce, with minimal environmental damage and external energy input in the form of nitrogenous and phosphorous fertilizers. In present studies group of plants were tested for their actual

position based on various morphological characters and genomic information. This study shows plants belongs to Fabaceae (Papilionaceae), Mimosaceae, Caesalpinaceae based on morphological characters has different members and the based on the genomic characteristics they fall in different groups. We conclude that as per botanical classification of species belonging to Leguminosae families are classified differently based on morphological character with compare to genomic characters.

Keywords: Leguminosae family, Bioinformatics, NCBI, rbcL.

1. INTRODUCTION

1.1. Leguminosae family

The Leguminosae family contains species of Plants, Herbs, Shrubs, and Trees. Legumes are used as crops, forages and green manures; they also synthesize a wide range of natural products such as flavours, drugs, poisons and dyes. Legumes are able to convert atmospheric nitrogen into nitrogenous compounds useful to plants. This is achieved by the presence of root nodules containing bacteria of the genus *Rhizobium*. These bacteria have a symbiotic relationship with Legumes, fixing free nitrogen for the plants; in return legumes supply the bacteria with a source of fixed carbon produced by photosynthesis. This enables many legumes to survive and compete effectively in nitrogen poor conditions. Leguminosae family is further classified into three subfamilies; Fabaceae (Papilionaceae), Caesalpinaceae and Mimosaceae.

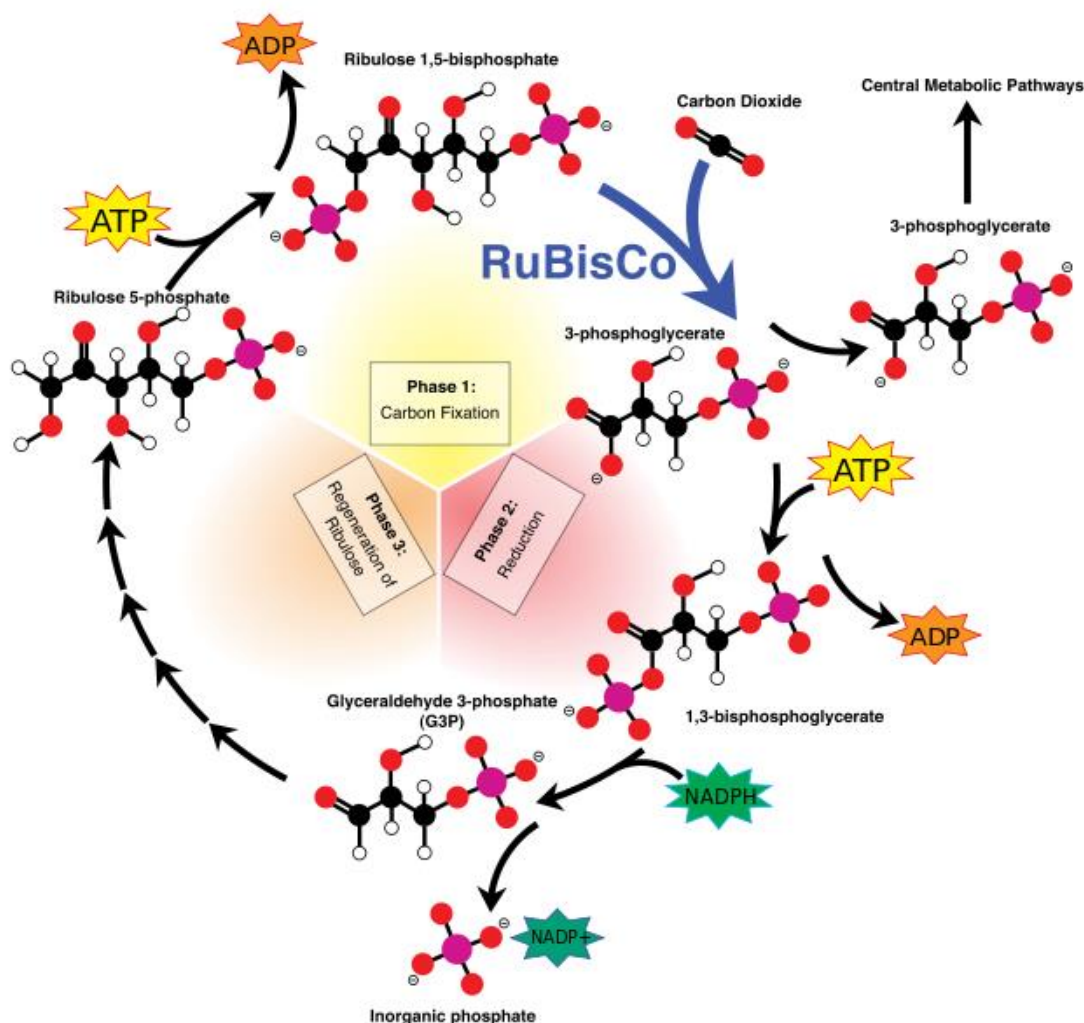


Figure 1

Example of rbcL gene which is involved in Carbon fixation in photosynthetic organisms

1.2. RbcL gene

The most common gene used for plant phylogenetic analyses is the plastid-encoded *rbcl* gene. This single copy gene is approximately 1430 base pairs in length and is free from length mutations except at the far 3' end. It has fairly conservative rate of evolution. The function of the *rbcl* gene is to code for the large subunit of ribulose 1, 5 biphosphate carboxylase/oxygenase (RUBISCO or RuBPCase). The sequence data of the *rbcl* gene are widely used in the reconstruction of phylogenies throughout the seed plants. Comprehensive phylogenetic analyses of Leguminosae began with the plastid gene *rbcl* following the early, widespread use of this gene for phylogenetic studies of land plant relationships.

- *rbcl* gene is involved in following pathways,
 1. Glyoxylate and dicarboxylate metabolism
 2. Carbon fixation in photosynthetic organisms
 3. Carbon metabolism
- During carbon fixation, the substrate molecules for RuBisCO are ribulose-1, 5-bisphosphate, and carbon dioxide (distinct from the "activating" carbon dioxide as shown in figure 1).

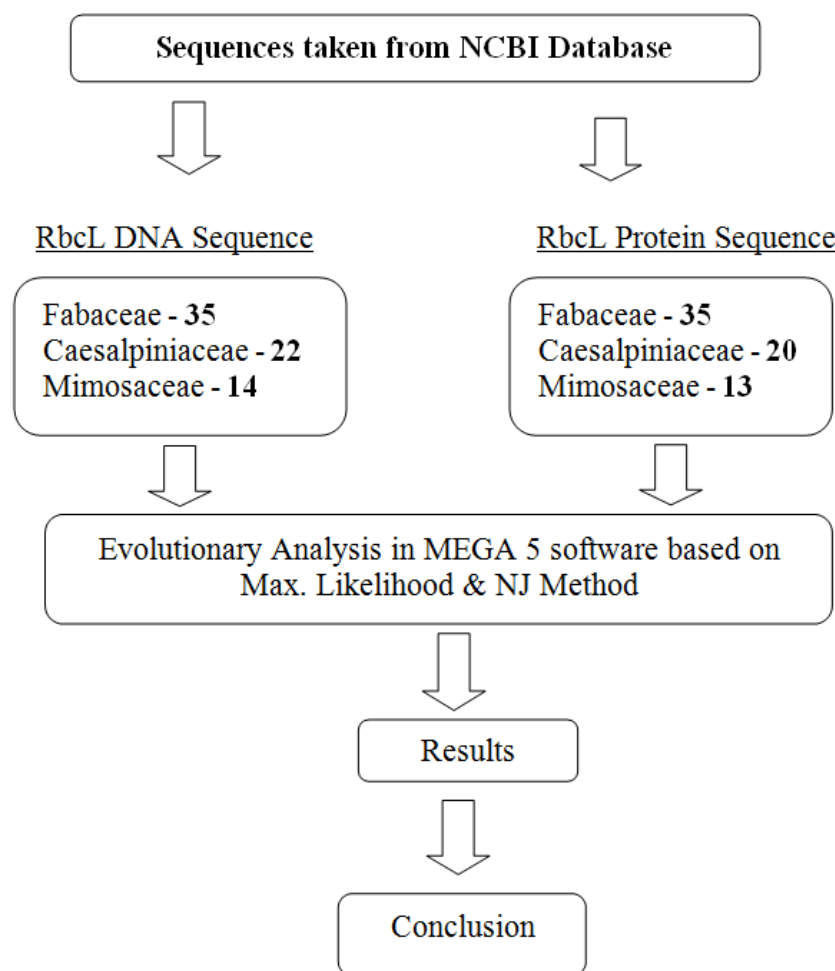


Figure 2

Flow chart of method

1.3. NCBI (The National Center for Biotechnology Information)

The National Center for Biotechnology Information (NCBI) is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health. The NCBI houses a series of databases relevant to biotechnology and biomedicine. Major databases include GenBank for DNA sequences, Protein, Genome, EST etc. All these databases are available online through the Entrez search engine. (<http://www.ncbi.nlm.nih.gov/>. The National Center for Biotechnology Information (NCBI 2001)).

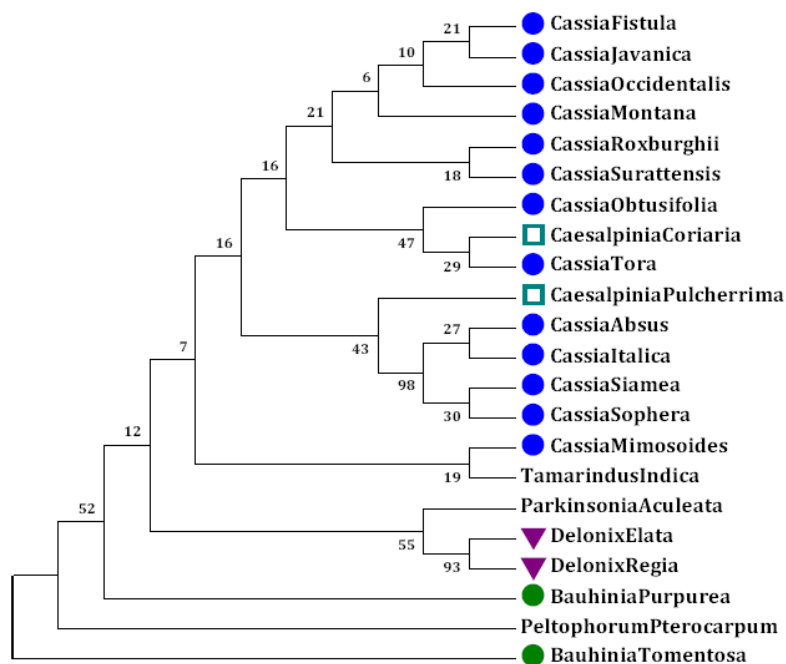


Figure 3

Result of Max.Likelihood Bootstrap Method

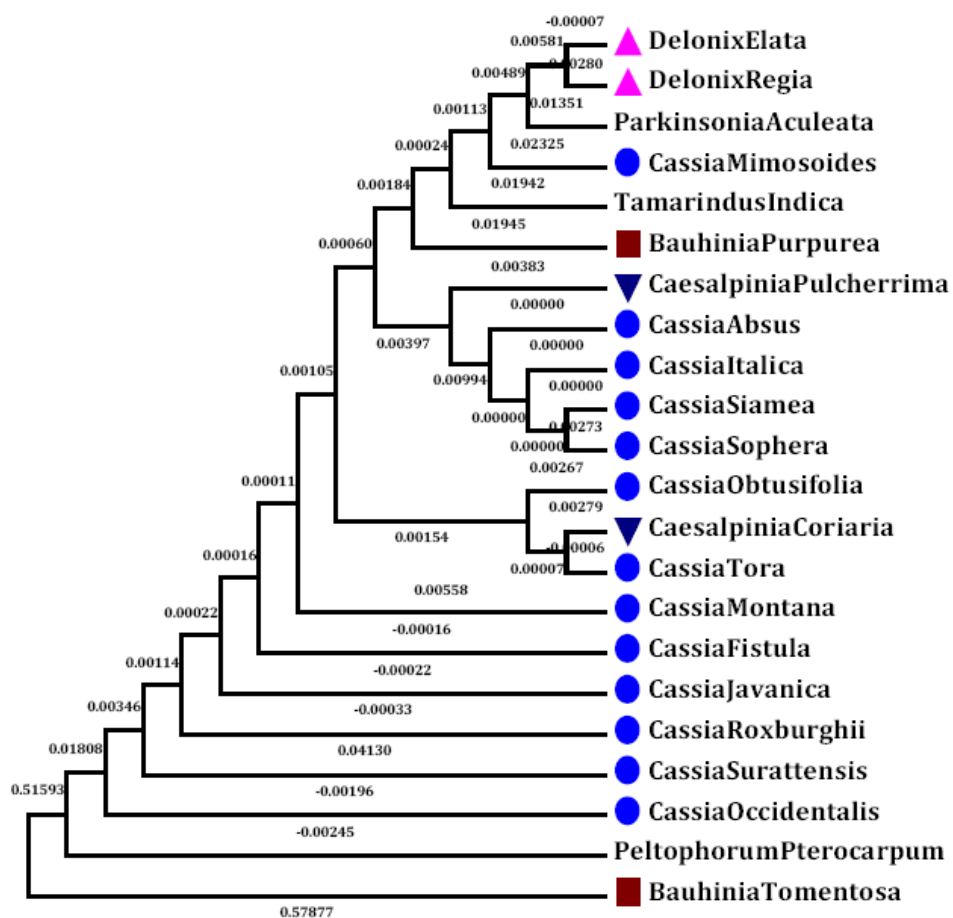


Figure 3.1

Result of NJ Method

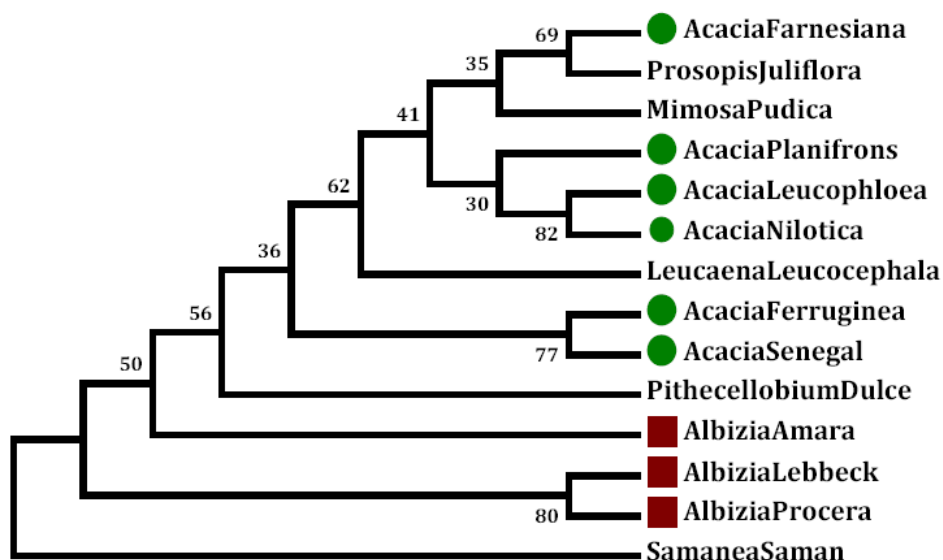


Figure 4

Result by Max.Likelihood Bootstrap Method

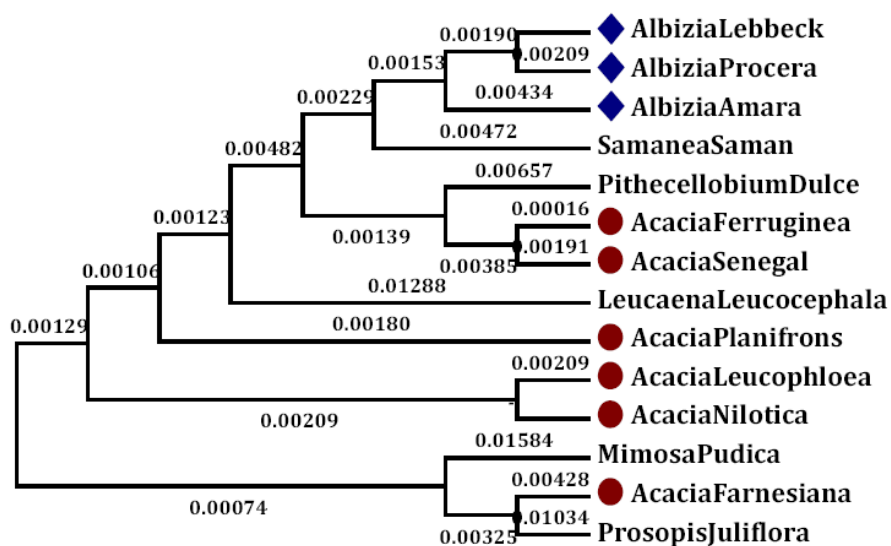


Figure 4.1

Result by NJ Method

1.4. DNA (Deoxyribonucleic acid) / Nucleotide

The Deoxyribonucleic acid (DNA) is a molecule that encodes the genetic instructions used in the development and functioning of all known living organisms and many viruses. Along with RNA and proteins, DNA is one of the three major macromolecules essential for all known forms of life. Genetic information is encoded as a sequence of nucleotides (guanine, adenine, thymine, and cytosine) recorded using the letters G, A, T, and C. Most DNA molecules are double-stranded helices, consisting of two long polymers of simple units called nucleotides, molecules with backbones made of alternating sugars (deoxyribose) and phosphate groups (related to phosphoric acid), with the nucleobases (G, A, T, C) attached to the sugars. DNA is well-suited for biological information storage, since the DNA backbone is resistant to cleavage and the double-stranded structure provides the molecule with a built-in duplicate of the encoded information.

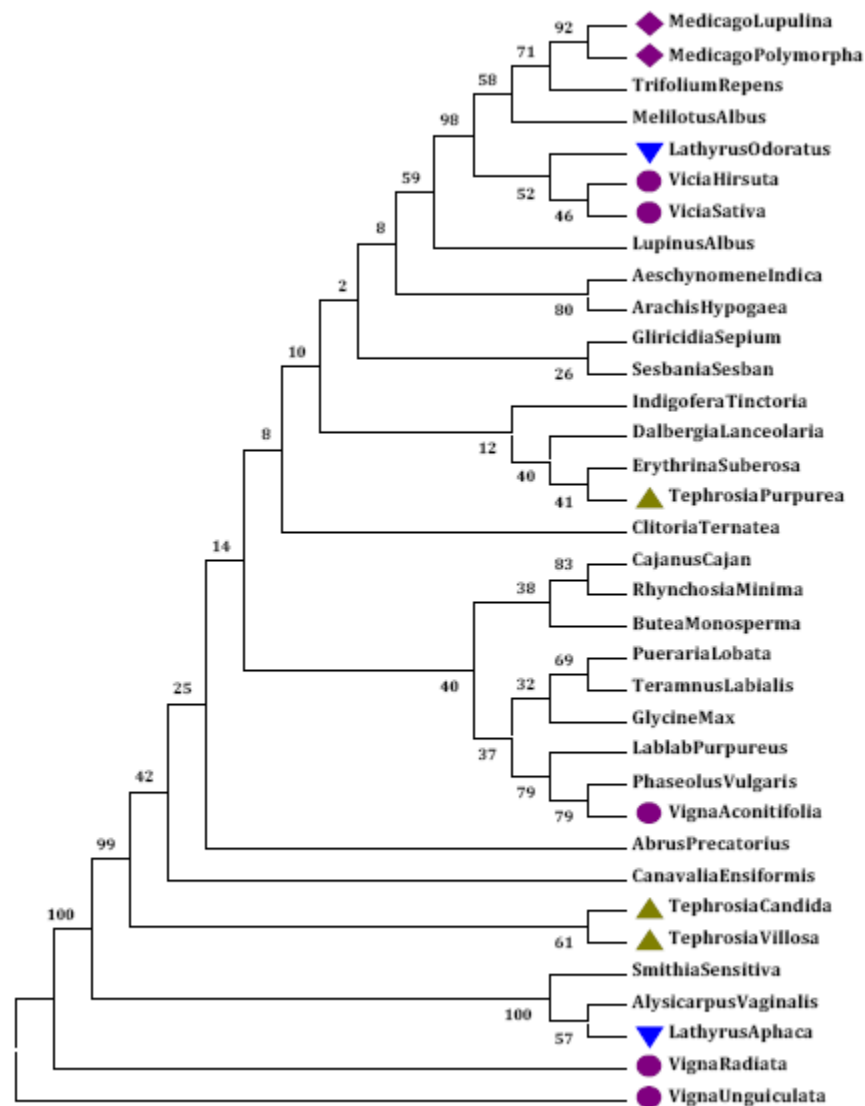


Figure 5

Result by Max.Likelihood Bootstrap Method

1.5. Protein

Proteins are large biological molecules consisting of one or more chains of amino acids. Proteins perform a vast array of functions within living organisms, including catalyzing metabolic reactions, replicating DNA, responding to stimuli, and transporting molecules from one location to another. Proteins differ from one another primarily in their sequence of amino acids, which is dictated by the nucleotide sequence of their genes, and which usually results in folding of the protein into a specific three-dimensional structure that determines its activity.

2. MATERIALS AND METHODS

In this paper we have considered around 266 species which are found in Gujarat state of India. Further we searched each species in NCBI database and finally found around 149 species' information like DNA, Protein and other useful information of leguminosae family. Further we have only considered rbcL gene sequences of DNA and Protein sequences. Analysis done in MEGA software and analysis was done with Maximum Likelihood Method with Bootstrap method and NJ method also (Figure 2).

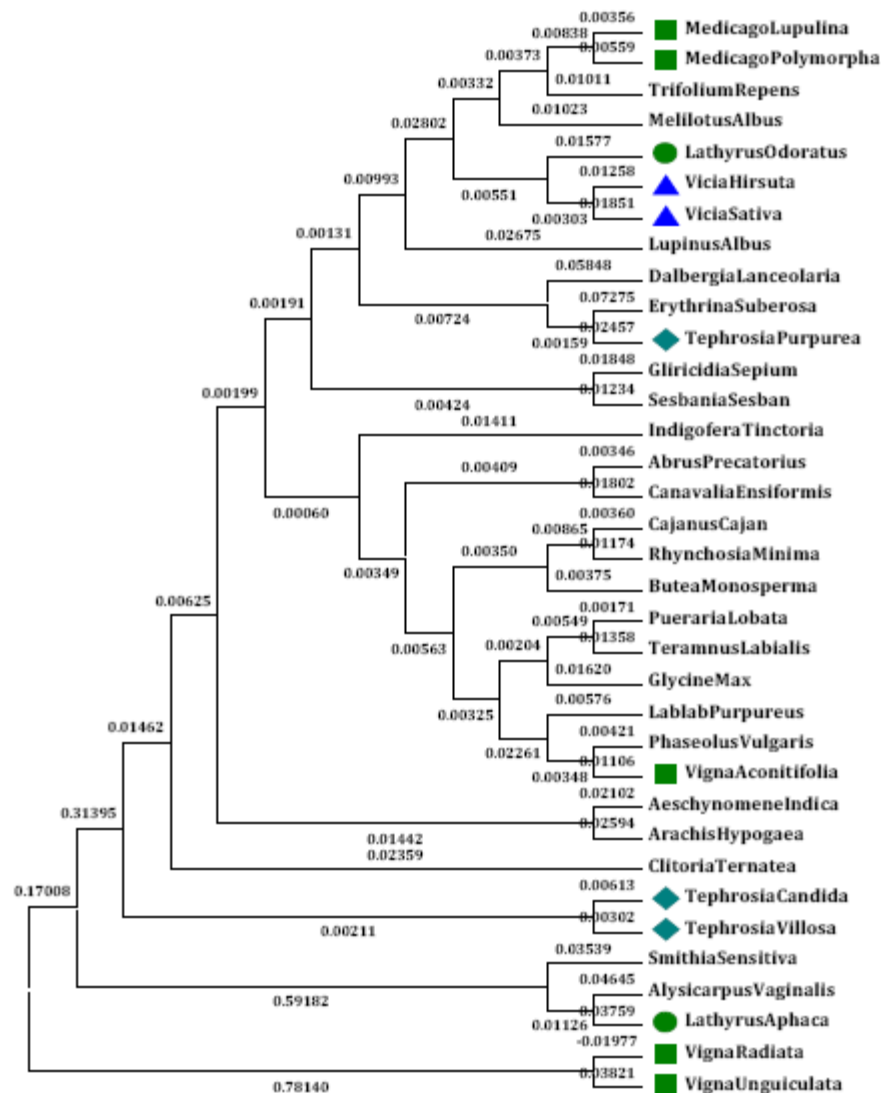


Figure 5.1

Result by NJ Method

3. RESULTS

3.1. Results of DNA *RBCL* gene sequences

3.1.1. Caesalpinieae subfamily

As shown in figure 3 which is result of Max.Likelihood method, species of *Cassia*, *Caesalpinia*, *Delonix* and *Bauhinia* genus are related as per morphological characters or botanical classifications except one species *Cassia tora* which is places between two species of *Caesalpinia* genus which are related by common node. Figure 3.1, which is result of NJ method, *Delonix* and *Cassia* genus are related as per morphological characters or botanical classifications but species of *Bauhinia* and *Caesalpinia* genus are distantly related with each other which is not true when we compare it with morphological characters or as per botanical classification. If we compare two figures 3 & 3.1 we observed that some species like *Peltophorum pterocarpum*, *Bauhinia tomentosa*, *Delonix elata*, *Delonix regia*, *Cassia fistula*, *Cassia javanica*, *Cassia roxburghii*, *Cassia surattensis*, *Cassia absus*, *Cassia italica*, *Cassia siamea* and *Cassia sophera*, there sequential order remain unchanged in both results. So, there might be strong sequence similarity between them so that they remain unchanged in two different results which are done with two different methods.

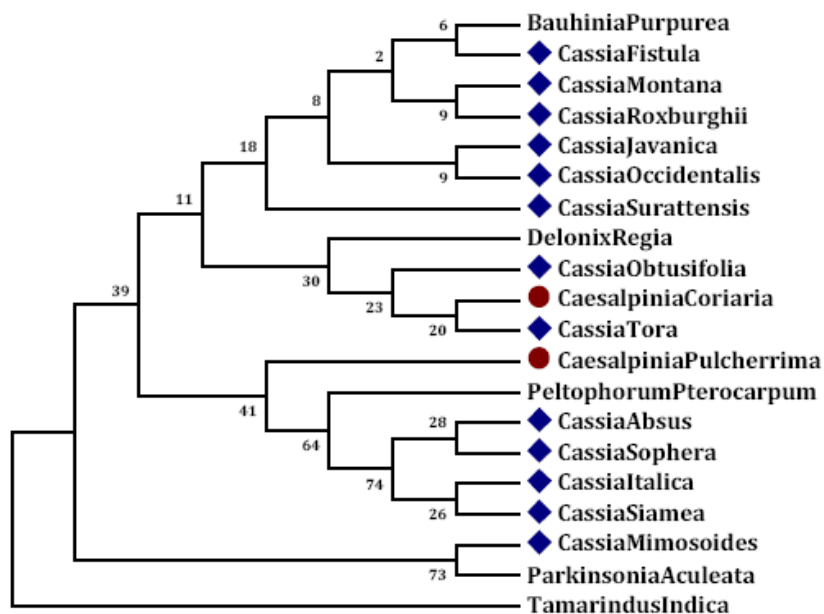


Figure 6

Result by Max.Likelihood Bootstrap Method

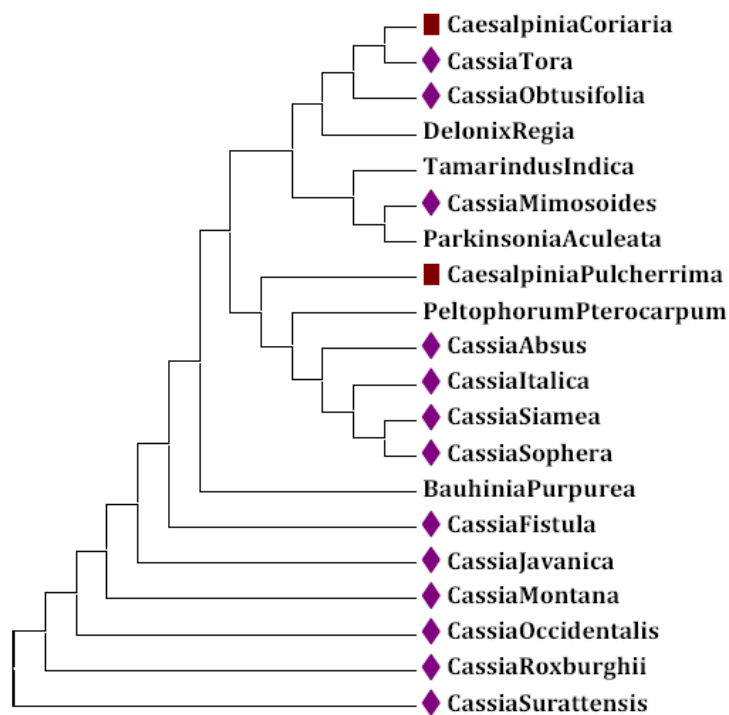


Figure 6.1

Result by NJ Method

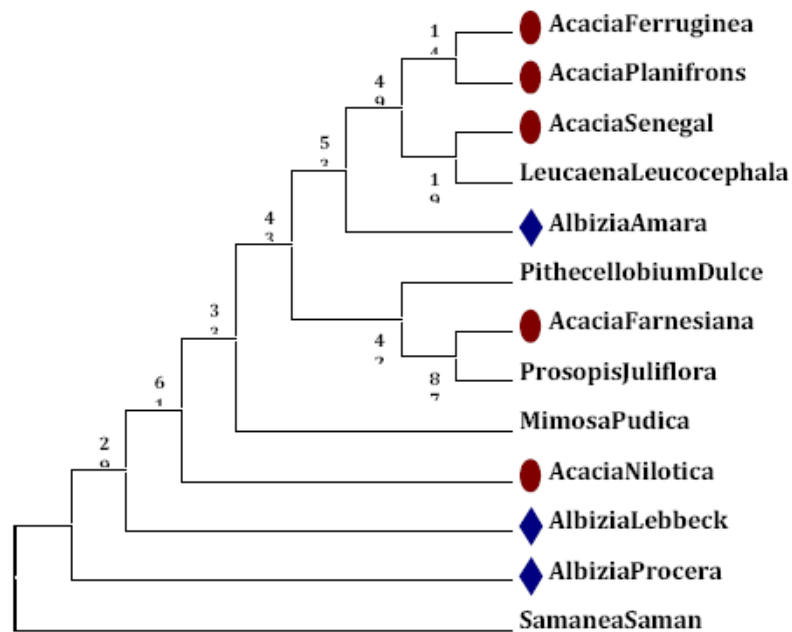


Figure 7

Result by Max.Likelihood Bootstrap Method

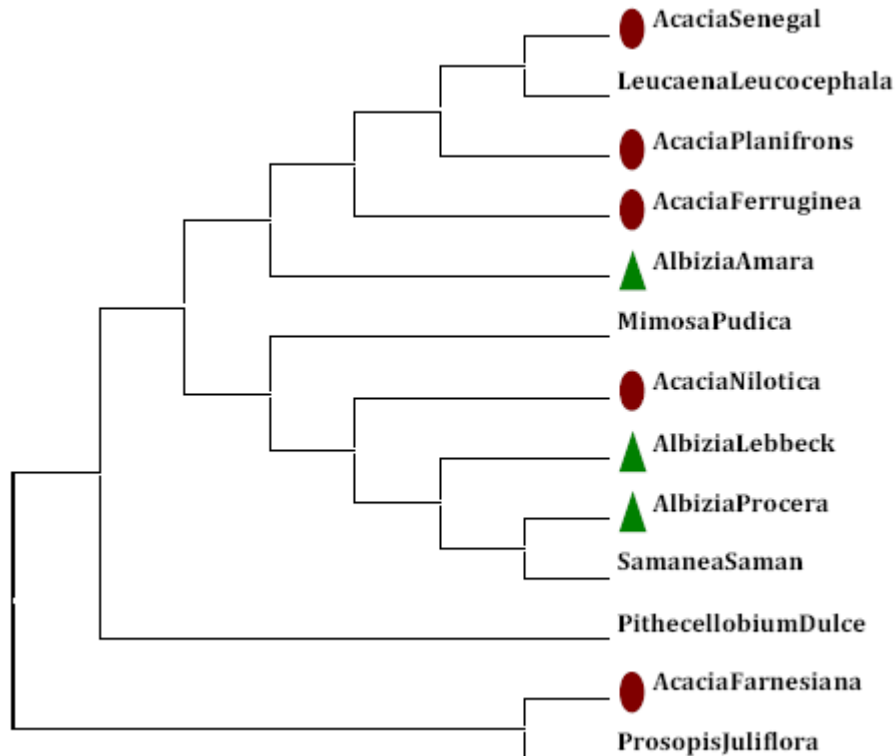


Figure 7.1

Result by NJ Method

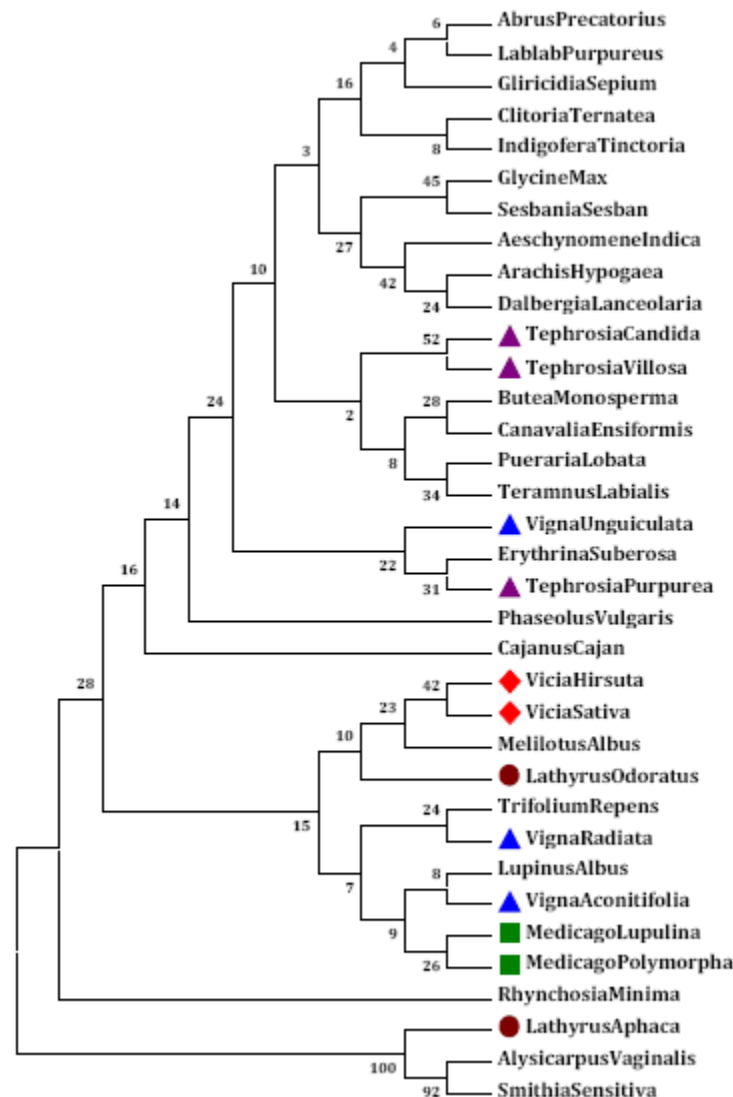


Figure 8

Result by Max.Likelihood Bootstrap Method

3.1.2. Mimosaceae subfamily

As shown in figure 4 which is result of Max.Likelihood method, species of genus *Albizia* and few species of *Acacia* genus are related as per morphological characters or botanical classifications except four different species which are placed between species of *Acacia* genus. Figure 4.1, which is result of NJ method, *Albizia* and *Acacia* genus are related as per morphological characters or botanical classifications but two species *Leucaena leucocephala* and *Mimosa pudica* fall within *Acacia* genus which may contain common DNA bases with other species of *Acacia*. If we compare two figures 4 & 4.1 we observed that some species like *Acacia planifrons*, *Acacia leucophloea*, *Acacia nilotica*, *Albizia lebbeck* and *Albizia procera*, their sequential order remain unchanged in both results. So, there might be strong sequence similarity between them that they remain unchanged in two different results which are done with two different methods.

3.1.3. Fabaceae family

As shown in figures 5 & 5.1 which are results of Max.Likelihood and NJ method, species of genus *Medicago*, *Vicia* and few species of *Vigna* and *Tephrosia* genus are related as per morphological characters or botanical classifications except species like *Tephrosia purpurea* and *Vigna aconitifolia* which are distantly related with other species of their genus. But species of *Lathyrus* genus are distantly related with each other. If we compare two figures 5 & 5.1 we observed that there are many species have not changed in

both results also their sequential order remain unchanged. So, there might be strong sequence similarity between species that they remain unchanged in two different results which are done with two different methods.

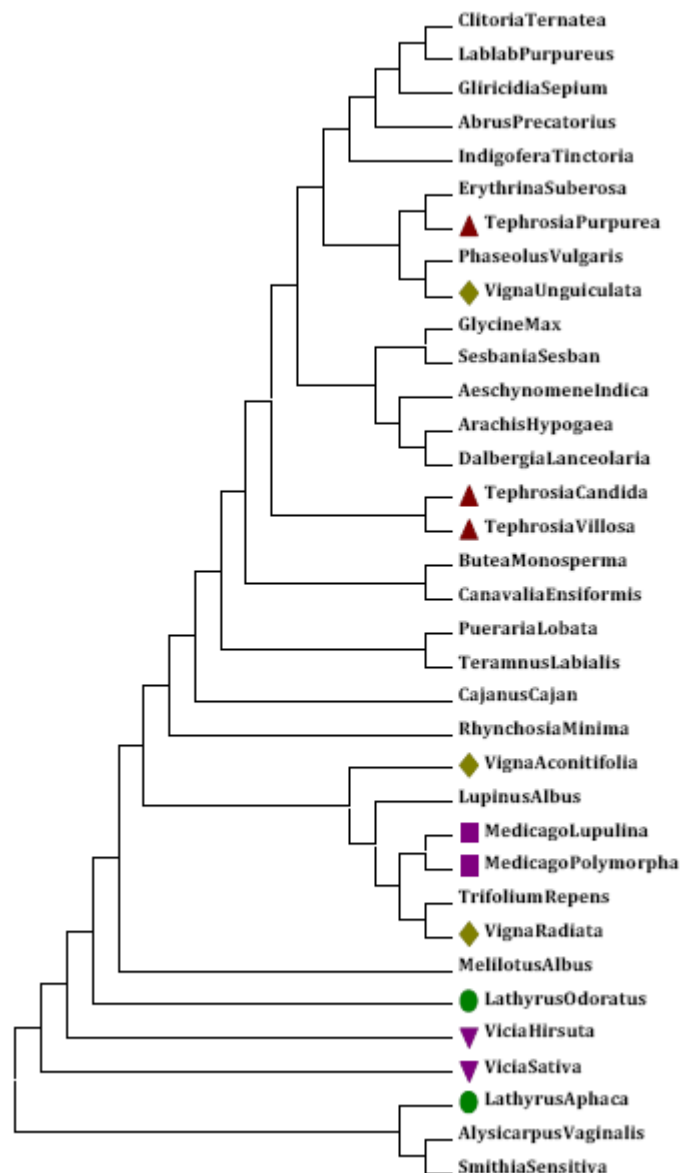


Figure 8.1

Result by NJ Method

3.2. Results of *PROTEIN RBCL* gene sequences

3.2.1. *Caesalpinieae* subfamily

As shown in figure 6 which is result of Max.Likelihood method, species of *Cassia* and *Caesalpinia* genus are related as per morphological characters or botanical classifications except one species *Cassia tora* which is placed between two species of *Caesalpinia* genus which are related by common node. *Delonix regia* and *Peltophorum pterocarpum* these two species are placed within *Cassia* genus, so these two species may have common base pairs which are closely related with their nearby species. Figure 6.1, which is result of NJ method, *Cassia* genus is related as per morphological characters or botanical classifications except few species also fall within *Cassia* genus. Species of *Caesalpinia* genus are distantly related with each other which are not true when we compare it with morphological characters or as per botanical classification.

3.2.2. Mimosaceae Subfamily

As shown in figure 7 which is result of Max.Likelihood method, species of genus *Albizia* and species of *Acacia* genus are related as per morphological characters or botanical classifications except *Acacia farnesiana* and *Acacia nilotica* which are not closely related with *Acacia* genus and *Albizia amara* species is not closely related with other species of *Albizia* species. Figure 7.1, which is result of NJ method, *Albizia* and *Acacia* genus are related as per morphological characters or botanical classifications *Acacia farnesiana* and *Acacia nilotica* which are not closely related with *Acacia* genus and *Mimosa pudica* fall within *Albizia* genus which may contain common amino acids with other species of *Albizia*. If we compare two figures 7 & 7.1 we observed that species like *Albizia lebbeck* and *Albizia procera*, there sequential order remain unchanged in both results. So, there might be strong sequence similarity between them that they remain unchanged in two different results which are done with two different methods.

3.2.3. Fabaceae family

As shown in above figures 8 & 8.1 which are results of Max.Likelihood and NJ method, species of genus *Medicago*, *Vicia* and few species of *Vigna* and *Tephrosia* genus are related as per morphological characters or botanical classifications except species like *Tephrosia purpurea* and *Vigna unguiculata* which are distantly related with other species of their genus. But species of *Lathyrus* genus are distantly related with each other. If we compare two figures 8 & 8.1 we observed that there are many species have not changed in both results also there sequential order remain unchanged. So, there might be strong sequence similarity between species that they remain unchanged in two different results which are done with two different methods.

4. CONCLUSION

Reconstructing the phylogenetic relationships of the Leguminosae family is essential for understanding the origin and diversification of this ecologically and economically important family of angiosperms. This study shows plants belongs to Fabaceae (Papilionaceae), Mimosaceae, Caesalpiniaceae based on morphological characters they are classified differently and they fall in different groups based on DNA and Protein sequences of *rbcl*. In this study we observed that as per the botanical classification species are classified differently based on their morphological features like species' flower color, size and shape, types and arrangements of Stipules, size of plant etc. But this study focus on evolutionary relationship of Leguminosae Family species based on DNA & Protein sequences of *rbcl* gene with Multiple sequence alignment by Maximum likelihood & Neighbor Joining method where we observed that some species belonging to same genus are fall very nearly as per botanical classification which is correct as per both botanical and evolutionary relationship but we observed that few species are distantly related even if they are from same genus. Even some species' position also remains unchanged in both Max.Likelihood method and in NJ method in each subfamily results.

ACKNOWLEDGMENT

We would like to thank Sardar Patel University and PURSE program of DST- SPU for their support and Department Of Biotechnology (DBT), Ministry of Science & Technology, Government of India.

REFERENCE

- Harborne JB. Phytochemistry of the Leguminosae. In Phytochemical Dictionary of the Leguminosae, eds Bisby, F.A. et al. London: Chapman & Hall, 1994
- Hayashi K, Kawano S. Molecular systematics of *Lilium* and allied genera (Liliaceae): phylogenetic relationships among *Lilium* and related genera based on the *rbcl* and *matK* gene sequence data. *Plant Species Biology* 2000, 15, 73-93
- Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evans R, Sauquet H, Neinhuis C, Slotta TAB, Jens GR, Campbell CS, Chatrou LW. Angiosperm phylogeny based on *matK* sequence information. *American Journal of Botany* 2003, 90, 1758-1776
- Hilu KW, Liang H. The *matK* gene: sequence variation and application in plant systematics. *American Journal of Botany* 1997, 84, 830-839
- <http://plantnet.rbgsyd.nsw.gov.au/iopi/iopihome.htm>
- <http://www.en.wikipedia.org>
- <http://www.ildis.org/>
- <http://www.kew.org/>
- <http://www.missouribotanicalgarden.org/>
- <http://www.ncbi.nlm.nih.gov/>
- <http://www.ncbi.nlm.nih.gov/nuccore/>
- <http://www.ncbi.nlm.nih.gov/protein/>
- <http://www.theplantlist.org/browse/A/Leguminosae>
- Jean-Michel Claverie and Cedric Notredame. Bioinformatics – A Beginner's Guide. Publ. by Wiley Publishing, Inc. USA, 2003
- Kalpesh Anjaria. Ph. D. Thesis: Floristic studies of Anand District. Submitted to Sardar Patel University, Vallabh Vidyanagar, Anand, India, 2002

16. Martin F Wojciechowski, Matt Lavin, Michael J Sanderson. A Phylogeny Of Legumes (Leguminosae) Based On Analysis Of The Plastid Matk Gene Resolves Many Well-supported Subclades Within The Family
17. Oza GM, Kishore S Rajput. Biodiversity of Gujarat Forest Trees. Publ. By INSONA, Vadodara, India, 2006
18. Patel, Anjaria, Panchal. Leguminous Trees In Anand District: Collection and Analysis With Bioinformatics Applications. LAP LAMBERT Academic Publishing, Germany, 2012
19. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 5; Tamura, Peterson, Stecher, Nei, Kumar, 2011
20. Polhill RM, Raven PH. (eds). Advances in Legume Systematics. *Royal Botanic Gardens, Kew*, 1981
21. Sagar Patel, Panchal H, Anjaria K. DNA Sequence analysis by ORF FINDER & GENOMATIX Tool: Bioinformatics Analysis of some tree species of Leguminosae Family. Publication Year: 2012, Page(s): 922- 926. Proceedings of the *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Philadelphia, USA, Oct 4-7, 2012; E-ISBN: 978-1-4673-2744-2; Print ISBN: 978-1-4673-2746-6
22. Sagar Patel, Panchal H, Smart J, Anjaria K. Species Information Retrieval Tool: A Bioinformatics tool for Leguminosae family in *International Journal of Bioinformatics and Biological Science*, 2013, 1(2), 187-194
23. Sagar Patel, Panchal H. Leguminobase: A Tool To Get Information Of Some Leguminosae Family Members From NCBI Database in *Journal of Advanced Bioinformatics Applications and Research*, 2013, 4(3), 54-59
24. Sagar Patel, Panchal H., Anjaria K., Phylogenetic analysis of some leguminous trees using CLUSTALW2 Bioinformatics Tool. Publication Year: 2012, Page(s): 917- 921. Proceedings of the *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Philadelphia, USA, Oct 4-7, 2012; E-ISBN: 978-1-4673-2744-2; Print ISBN: 978-1-4673-2746-6
25. Sagar Patel, Panchal H., Smart J., Anjaria K., 2013. Distribution of Leguminosae family members in Gujarat State of India: Bioinformatics Approach in *International Journal of Computer Science and Management Research*, 2013, 2(4), 2184-2189
26. Sagar Patel. Some Leguminous trees in Anand District (M.Sc., Project work) Sardar Patel University, Vallabh Vidyanagar, Gujarat, India, 2011
27. Shah GL. Flora of Gujarat State. Publ. by Sardar Patel University, Vallabh Vidyanagar, Anand, India, 1978
28. Smartt J, Simmonds NW. (eds). Evolution of Crop Plants. *Harlow: Longman Scientific & Technical*, 1995
29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* (submitted), 2011